Biomedical Optics EXPRESS

Label-free multimodal imaging with simultaneous two-photon and three-photon microscopy and kernel-based nonlinear scaling denoising

WENTAO WU,[1,2] CHRISTOPH BRANDT,[1] XIN ZHOU,[1] AND SHUO TANG[1,3]

[1] *Department of Electrical and Computer Engineering, University of British Columbia, 5500-2332 Main Mall, Vancouver, BC V6T 1Z4, Canada*
[2] *wuwentao@protonmail.ch*
[3] *tang@ece.ubc.ca*

Abstract: We report on a compact multimodal imaging system that can acquire two-photon microscopy (2PM) and three-photon microscopy (3PM) images simultaneously. With dual excitation wavelengths, multiple contrasts including two-photon-excitation-fluorescence (2PEF), second harmonic generation (SHG), and third harmonic generation (THG) are acquired simultaneously from cells, collagen fibers, and interfaces, all label-free. Challenges related to the excitation by two wavelengths and the effective separation of 2PM and 3PM signals are discussed and addressed. The data processing challenge where multiple contrasts can have significantly varying signal levels is also addressed. A kernel-based nonlinear scaling (KNS) denoising method is introduced to reduce noise from ultra-low signal images and generate high-quality multimodal images. Simultaneous 2PM and 3PM imaging is demonstrated on various tissue samples. The simultaneous acquisition speeds up the imaging process and minimizes the commonly encountered problem of motion artifacts and mechanical drift in sequential acquisition. Multimodal imaging with simultaneous 2PM and 3PM will have great potential for label-free *in-vivo* imaging of biological tissues.

© 2023 Optica Publishing Group under the terms of the Optica Open Access Publishing Agreement

1. Introduction

Multiphoton microscopy (MPM) is a powerful and widely utilized technique for non-invasive and label-free imaging of biological tissues [1–8]. Multiple contrasts can be acquired by MPM using different nonlinear optical effects. For example, in two-photon microscopy (2PM), two-photon-excitation-fluorescence (2PEF) can be acquired from intrinsic fluorophores such as nicotinamide adenine dinucleotide hydrogen (NADH) in cells, and second harmonic generation (SHG) can be acquired from non-central symmetric molecules such as collagen fibers; in three-photon microscopy (3PM), three-photon-excitation-fluorescence (3PEF) can also be acquired from intrinsic fluorophores, and third harmonic generation (THG) can be acquired from interfaces such as the lipid aqueous medium interface [9–15]. Being able to acquire multiple contrasts is important in label-free imaging as it can provide a comprehensive understanding of tissue morphology and function. With the advancement in lasers as well as signal detection techniques, combining 2PM and 3PM in a single system has gained a lot of interest as it can provide complementary information about tissues including cells, collagen fibers, lipids, etc., and form information-rich images that are essential for label-free tissue imaging [16–21].

Several groups have demonstrated combined 2PM and 3PM imaging systems. F. Akhoundi et al. demonstrated a 2PM (SHG) and 3PM (3PEF and THG) probe using a single excitation wavelength at 1700nm [22]. However, their THG and 3PEF signals were acquired sequentially by

#504550
Journal © 2024

https://doi.org/10.1364/BOE.504550
Received 1 Sep 2023; revised 1 Dec 2023; accepted 1 Dec 2023; published 6 Dec 2023

a single photomultiplier tube (PMT), and samples were hematoxylin and eosin (H&E) stained for 3PEF imaging. At 1700-nm excitation wavelength, it is not possible to excite 2PEF from tissues without staining, and 3PEF is a higher-order nonlinear effect with much lower efficiency than 2PEF. Although excitation by a single wavelength is simple, it has many limitations in combined 2PM and 3PM imaging. On one hand, lasers with long wavelengths (>1 µm) cannot efficiently excite 2PEF from intrinsic fluorophores in tissues [23,24]. On the other hand, a short excitation wavelength (<1 µm) is not optimal for THG imaging as the signal wavelength will fall inside the ultraviolet regime [23].

To achieve both 2PM and 3PM from tissues label-free, several other groups including our group have used two excitation wavelengths, where a shorter wavelength (<1 µm) is for 2PM and a longer wavelength (>1 µm) for 3PM, respectively [16,17,24–27]. D. Septier et al. demonstrated 2PM (SHG) and 3PM (3PEF and THG) in mouse skin excited by 920 and 1300 nm, respectively [26]. By using a single PMT, all three images were acquired one after another. F. Aptel et al. performed 2PM (2PEF and SHG) and 3PM (THG) imaging of human cornea using dual excitation wavelengths [27], where a 730-nm wavelength excited 2PEF and a 1200-nm wavelength excited SHG and THG. The two wavelengths were applied on tissue in sequential mode. Previously, our group has demonstrated combined 2PM (2PEF and SHG) and 3PM (THG) using dual excitation wavelengths at 790 and 1580 nm obtained from a frequency-doubled erbium-doped fiber laser [16,17,28]. 3D volumetric 2PM and 3PM image stacks were acquired in sequential mode, where the 790- and 1580-nm excitation wavelengths were applied to tissue one after another. There are several drawbacks to the sequential acquisition of 2PM and 3PM images. First, it doubles the image acquisition time because the 790- and 1580-nm excitation wavelengths are applied one after another. Second, it is vulnerable to motion artifacts and mechanical drifts which can cause spatial mismatch when merging the 2PM and 3PM images. The acquisition of corresponding frames in 2PM and 3PM stacks are separated by the duration of imaging a whole stack which can last several minutes depending on the number of frames in a stack [16]. Meanwhile, motion artifacts are especially critical during *in-vivo* imaging.

To reduce motion artifacts and mechanical drifts in the sequential acquisition of 2PM (2PEF and SHG) and 3PM (THG) imaging, A. Filippi et al. used computer-controlled mechanical shutters to alternate wavelengths for 2PM and 3PM between consecutive frame acquisition [29]. The 2PEF and THG images were excited by wavelengths of 800 and 1200 nm, respectively, and the SHG images were excited by both wavelengths [29]. The frame-level interleaving mitigated inconsistency caused by sample motion or mechanical drift between 2PM and 3PM images. However, for each frame (depth) in a stack, the 2PM and 3PM images were still acquired sequentially, and the total imaging time would still be doubled compared with simultaneous acquisition.

Temporal multiplexing can further improve the simultaneous acquisition of 2PM and 3PM signals at the pixel level by interleaving 2PM and 3PM excitation pulses [30,31]. D. G. Ouzounov et al. interleaved 920-nm pulses from a Ti:sapphire laser modulated with an electro-optical modulator with 1300-nm pulses from a noncollinear optical parametric amplifier (400-kHz repetition rate) in 2.5-µs periods, respectively [30]. They performed simultaneous recording of neuron activity with 2PEF and 3PEF on GCaMP6-labeled neurons at the same site and time in order to validate functional imaging of 3PM by comparing it with 2PM [30]. The temporal multiplexing approach allowed the simultaneous acquisition of 2PM and 3PM signals within µs time delay. In order to increase the imaging volume in neuroscience study, S. Weisenburger et al. utilized spatiotemporal multiplexing of 2PM and 3PM to image multiple sub-volumes at the same time [31]. A custom laser system outputted 960 nm for 2PM and 1300 nm for 3PM excitation, at repetition rates of 4.7 and 1.0 MHz, respectively. They interleaved four 960-nm pulses and one 1300-nm pulse, each delayed by 8 ns through optical delay lines, forming five beamlets, each targeting a different sub-volume. Specifically, the 1300-nm beamlet was used to target a deeper axial location due to its higher penetration depth. Temporal multiplexing provides an alternative approach for signals that cannot be separated in the spectral domain. In both Refs. [30,31], 2PEF and 3PEF signals were acquired from GCaMP6-labeled neurons where the fluorescence emission signals could not be separately spectrally. However, temporal multiplexing generally requires high system complexity due to the requirements on modulators, optical delay lines, precise timing, and synchronization with the detection unit.

In this paper, we will address the challenges of applying two excitation wavelengths on tissue simultaneously for multimodal 2PM and 3PM imaging. We are aiming for a compact fiber-based system design and all label-free imaging. When two excitation wavelengths are shone on the sample simultaneously, the separation of 2PM and 3PM signals becomes difficult due to a partial wavelength overlapping between 2PEF and THG. Other challenges include the risk of thermal damage due to the additional excitation beam, and the difficulty to align two beams accurately for imaging the same field of view if two separate laser sources are utilized.

In addition to the challenges in the hardware design for data acquisition, the data processing of combined 2PM and 3PM imaging also has its unique challenges. While the different contrasts (e.g. 2PEF, SHG, and THG) visualize complementary tissue information, the signal level for each contrast can vary significantly, giving different detection channels highly differentiated signal-to-noise ratios (SNRs). This renders it difficult to visualize weak-signal channels, especially in merged multimodal images. For example, 3PM signals can be orders of magnitude weaker than 2PM signals [32]. Meanwhile, THG images from 3PM can reveal important information on tissue structures that lack 2PM contrasts, such as aqueous-lipidic interfaces in animal tissues and spores in plant tissues [16,20,33]. As another example, certain samples may only generate very weak (but still meaningful) signals on certain contrasts due to their structural and/or biochemical properties [16]. To reduce the amount of noise that would be introduced to the merged multimodal MPM images, images from the low-signal channels require denoising and SNR improvement before being merged with other channels.

Traditional denoising methods such as outlier removal and median filters are usually not effective in removing noise from those low-signal images while preserving tissue feature patterns [34]. Machine learning methods have also been utilized to denoise MPM images. Niu et al. demonstrated a neural-network-based method that effectively reduced noise from THG and 3PEF images of Hoechst blue labeled RAW 264.7 cells [35]. However, the feasibility of such methods typically relies on the availability of abundant training data [36]. In addition, the effectiveness of the trained neural network model may typically be limited to images with signal patterns similar to the training data. In recent years, self-supervised machine-learning methods have been developed to enable training using datasets without ground truth [37]. Noise2Void (N2 V) is a widely used self-supervised denoising algorithm that allows training using a single image without ground truth [38,39]. Nevertheless, machine-learning methods have the well-known black box problem, meaning that the trained model is typically not interpretable, which renders it difficult to explain how the image is processed [40].

In this paper, we will demonstrate a compact fiber-based multimodal MPM system with the capability of simultaneously acquiring 2PM (2PEF and SHG) and 3PM (THG) images. First, we will address the challenges related to data acquisition of simultaneous excitation using two wavelengths and the effective separation of 2PM and 3PM signals. Second, we will address the data processing challenges related to multimodal imaging. Specifically, a kernel-based nonlinear scaling (KNS) denoise and contrast enhancement method is introduced and demonstrated for improving images from low-signal channels and forming clean multimodal images. Comparisons are conducted between our denoising method and several other approaches including common traditional methods and N2 V. Finally, combined 2PM and 3PM imaging which simultaneously acquires 2PEF, SHG, and THG contrasts is demonstrated on biological samples, and the multimodal images are processed by our proposed denoising method to show the potential of this compact multimodal MPM system.

2. System setup

To perform simultaneous 2PM (2PEF and SHG) and 3PM (THG) imaging, a dual-wavelength excitation approach is used, where excitation wavelengths of ~790 nm and ~1580 nm are obtained from a frequency-doubled erbium-doped fiber laser with a pulse width of ~80 fs and a repetition rate at 50 MHz. Our laser system enables fiber-delivery of femtosecond pulses, which brings about advantages in compactness and portability for a potentially handheld imaging head. With the frequency-doubling scheme, both 2PM and 3PM can be achieved with the single laser source and that further enhances the compactness.

Figure 1 shows the optical layout of the combined 2PM and 3PM imaging system. The 1580-nm laser is partially frequency-doubled to 790 nm by a periodically poled MgO:LiNbO$_3$ (PPLN) crystal (MSHG1550-0.5-0.3, Covesion). The maximum conversion efficiency for the PPLN is 35%. However, as THG can be 10 times weaker than 2PM signals [32], the PPLN is placed slightly out of focus to preserve a larger portion of 1580 power for THG excitation. Thus, the conversion efficiency is controlled at ~18%. The laser power applied on samples is ~9 and ~40 mW at wavelengths of 790 and 1580 nm, respectively. The 790- and 1580-nm wavelengths are used to excite 2PM and 3PM signals, respectively. Compared with using two separate laser sources, the frequency-doubling solution avoids the difficulty of accurately aligning two beams since the fundamental and frequency-doubled beams are naturally overlapped. A miniaturized micro-electromechanical system (MEMS) mirror (13Z2.1-2400, Mirrorcle Technologies) performs XY scanning of the excitation beams. The pixel dwell time is set at 10 µs and that gives a frame rate of ~0.4 FPS. The miniaturized objective is custom-designed and consists of an aspheric lens (352080, LightPath) and a plano-convex lens (45469, Edmund Optics) [16,41]. The objective has a focal length of $f_{\mathrm{OBJ}} = 3.6$ mm and a numeric aperture (NA) of ~0.5. The aspheric lens is mounted inside a shape memory alloy (SMA) actuator which enables axial scanning. Depth-resolved image stacks are acquired by scanning the aspheric lens in the depth direction (z). The lateral and depth resolutions for 2PM imaging are ~0.7 and ~9 µm, respectively; and those for 3PM imaging are ~1 and ~15 µm, respectively.

The emitted 2PM and 3PM signals from the sample are collected backward through the same objective and are separated from the excitation beams by a long-pass dichroic beam splitter DBS1 (FF665-Di02, Semrock) with an edge wavelength of 665 nm. Signals are then coupled into a multimode fiber and delivered to the detection unit. Before being detected by PMTs, the multimodal signals are separated by two additional dichroic beam splitters DBS2 and DBS3. DBS2 is a long-pass dichroic beam splitter (FF414-Di01, Semrock) with an edge wavelength of 414 nm. SHG signal is reflected by DBS2 and then detected by PMT1 (H6780-20, Hamamatsu). The longer-wavelength lights containing 2PEF and THG signals are further separated by a notch dichroic beam splitter DBS3 (ZT532dcrb, Chroma). DBS3 has a central reflection wavelength of 532 nm and a full-width-half-maximum (FWHM) notch bandwidth of 25 nm. THG signal is reflected by DBS3 and fed into PMT3 (H7422, Hamamatsu). As shown in the spectral relationship in Fig. 1, due to a partial spectral overlapping between 2PEF and THG, a small portion of the 2PEF signal within the 532/25 nm band is also fed into PMT3. Thus, PMT3 contains THG + 2PEF signals. Most part of the 2PEF signal passes through DBS3 and then reaches PMT2 (H9305-03, Hamamatsu). The spectral relationship of the two excitation wavelengths and the three nonlinear signals of 2PEF, SHG, and THG are illustrated in Fig. 1. A spectrograph (SP-2150i, Princeton Instruments) and an EMCCD camera (ProEM 512, Princeton Instruments) are used to measure and verify the spectrum of the emitted signals.

The multimodal MPM system in Fig. 1 contains multiple considerations for a compact design, including the fiber-delivered excitation laser and fiber-based signal collection, the miniaturized

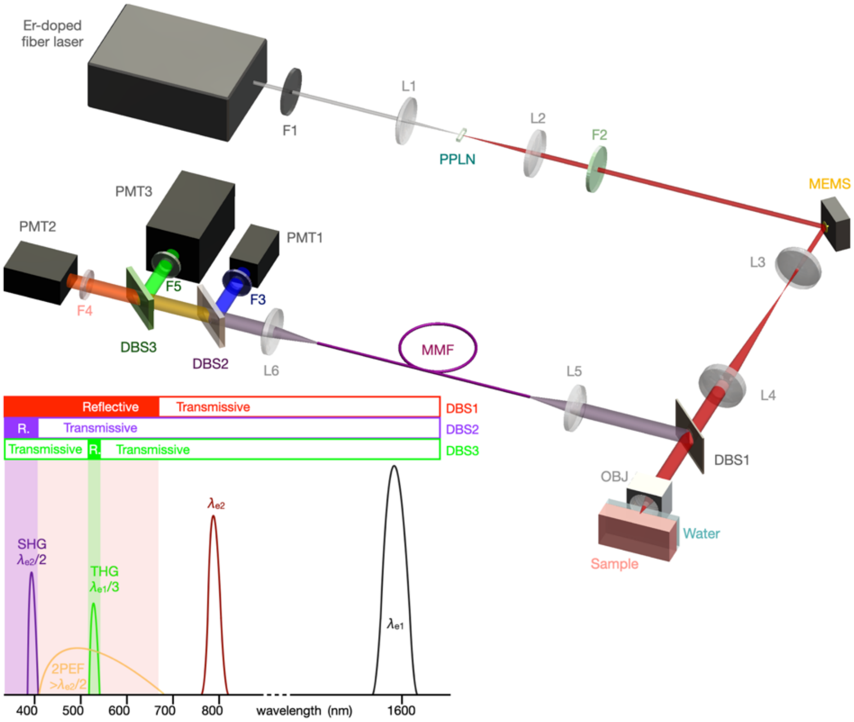
Er-doped fiber laser
F1
L1
PPLN
L2
F2
MEMS
L3
L4
DBS1
OBJ
Water
Sample
L5
MMF
L6
DBS2
F3
PMT1
PMT3
F5
DBS3
F4
PMT2
Reflective
Transmissive
DBS1
R.
Transmissive
DBS2
Transmissive
R.
Transmissive
DBS3
SHG
$\lambda_{e2}/2$
THG
$\lambda_{e1}/3$
2PEF
$>\lambda_{e2}/2$
λ_{e2}
λ_{e1}
400
500
600
700
800
wavelength (nm)
1600

Fig. 1. Optical layout of the compact multimodal MPM imaging system with the illustration of the spectral relationship of the excitation and emission signals. L – lens; PPLN – periodically poled MgO:LiNbO$_3$; MEMS – micro-electromechanical system; OBJ – objective lens pack; MMF – multimode fiber; DBS – dichroic beam splitter; PMT – photomultiplier tube; λ_e – excitation wavelength.

MEMS scanner and SMA actuator for XY and Z scanning, respectively, and the miniaturized objective with custom two-lens design. In the following sections, *ex-vivo* samples are imaged to demonstrate the proof of principle. Nevertheless, the compact design of this system will have the potential for *in-vivo* imaging.

3. Simultaneous acquisition of 2PM and 3PM signals

One challenge in the simultaneous dual-wavelength excitation comes from chromatic aberration of the objective lens between the two excitation wavelengths. For a compact design, the objective consists of a miniature aspheric lens and a plano-convex lens. While this lens design minimizes spherical aberration, there is still obvious chromatic aberration between the wavelengths of 790 and 1580 nm. Specifically, the 790- and 1580-nm wavelengths are focused to two different focal depths with a focal shift of ~80 µm as shown in Fig. 2(a). The amount of focal shift is lens-specific, and it is calculated by Zemax simulation and verified by experimental calibration using a silicon photonic chip that contains structures with a thickness of hundreds of nanometers. Therefore, the acquired 2PM and 3PM image stacks would have a ~80 µm depth offset, which can be compensated in post-processing by shifting the THG image stack downward by 80 µm. Figure 2(b) shows the verification of 2PM and 3PM signal spectra using the photonic chip sample. When the 790 nm is focused on the photonic chip, the SHG peak and the broadband 2PEF emission are detected in the spectrum; when the 1580 nm is focused on the photonic chip, the THG peak is observed in the spectrum.

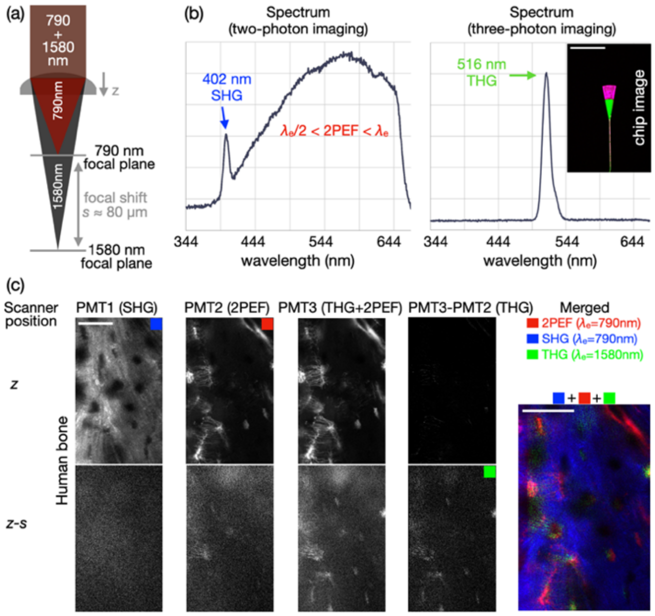

Fig. 2. Acquisition and verification of the multimodal MPM with simultaneous 2PM and 3PM imaging. (a) Focal shift between the two excitation wavelengths; (b) 2PM and 3PM signal spectra; (c) process for crosstalk removal and adjusting the depths between 2PM and 3PM stacks. The merged image is the superposition of single-channel images with the blue, red, and green marks, respectively. THG image is denoised before merging. Red – 2PEF; blue – SHG; green – THG. Scale bar is 40 µm.

Another challenge in simultaneous 2PM and 3PM imaging is the separation and simultaneous detection of multimodal signals. Based on the spectral relationship of the two excitation wavelengths and the 2PEF, SHG, and THG signals, dichroic beam splitters are carefully selected to spectrally separate the 2PEF, SHG, and THG signals. However, due to the partial spectral overlap between 2PEF and THG signals, PMT1, PMT2, and PMT3 detect SHG, 2PEF, and THG + 2PEF, respectively. Samples with both 2PEF and THG signals presenting in a considerable depth range may cause crosstalk from the 2PEF channel to the THG channel. Thin samples like the photonic chip are not significantly affected by this crosstalk issue. However, as shown in Fig. 2(c), the problem starts to appear when the sample thickness and imaging penetration reaches the focal shift of ~80 µm. The relatively low depth resolution of the miniature objective may cause the excitation of out-of-focus signals which can add difficulty to the separation of the THG signal. To obtain a clean THG image stack, an image post-processing method is proposed to remove the leaked 2PEF signal from THG images by performing a subtraction of $I_{\mathrm{THG}} = I_{\mathrm{PMT3}} - r \times I_{\mathrm{PMT2}}$ [42]. Here I_{THG} is the signal intensity of the THG image, and I_{PMT3} and I_{PMT2} stand for the signal intensity received by PMT3 and PMT2, respectively. In the equation of subtraction, the intensity of PMT2 is multiplied by a correction factor r because the two channels use different PMTs with different gains and the leaked portion of 2PEF signals can vary with sample.

The simultaneous acquisition of 2PM and 3PM signals doubles the imaging speed compared to sequential acquisition. Therefore, it can effectively reduce motion artifacts and mechanical drift between the 2PM and 3PM stacks. Figures 3(a) and 3(b) show the superposed SHG and THG images of a mouse femur bone acquired sequentially and simultaneously, respectively. The images are from stacks with 90 frames and acquired at a frame rate of 0.4 FPS. In the sequential acquisition, a relatively long time delay of ~5 minutes exists between the 2PM and 3PM images from the same depth. The delay includes time for stack acquisition, manual wavelength switch, and saving the files. This long separation in acquisition time makes the imaging vulnerable to mechanical drift. In the current design of our compact MPM system, the sample is mounted vertically during imaging, and immersion water is added between the objective and the sample. Gravity pulling on the sample and immersion water can cause mechanical drift. Mechanical drift can also originate from tension and elastic deformation in tissues. In Fig. 3, the SHG contrast comes from collagen fibers while the locations of osteocytes lack SHG signal and appear dark. The oval-shaped structures showing THG signals correspond to osteocytes which should fit into the dark holes shown in the SHG image. However, as can be seen in Fig. 3(a), the mechanical drift of the sample during the time between imaging the corresponding frames in 2PM and 3PM stacks (~5 min) caused motion artifacts on the superposed image. The location for the same osteocyte shifted ~6 µm between the SHG and THG images. This motion artifact is minimized by the simultaneous acquisition of 2PM and 3PM stacks and the misalignment between SHG and THG images is significantly reduced as shown in Fig. 3(b).

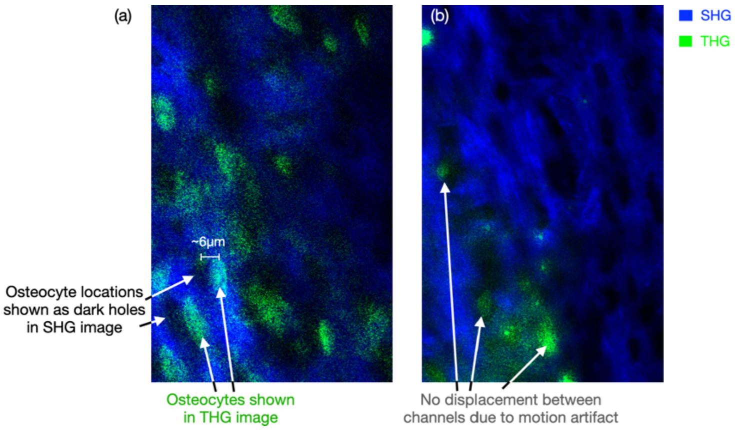
(a)
(b)
🟦 SHG
🟩 THG
~6µm
Osteocyte locations shown as dark holes in SHG image
Osteocytes shown in THG image
No displacement between channels due to motion artifact

Fig. 3. Motion artifacts comparison between (a) sequential and (b) simultaneous acquisition of 2PM and 3PM images of a mouse femur.

While the chromatic aberration of the miniaturized objective causes the scanning focal shift of ~80 µm and makes adjustment of stack images necessary when creating superposed images, the simultaneous scanning of two planes can potentially provide advantages in reducing thermal damage during multimodal imaging. To illustrate this, we have performed simulations of laser heating on tissue with and without the focal shift. The simulation is carried out using Ansys Workbench (Ansys Inc.) and the results are shown in Fig. 4. A volume of $200 \times 100 \times 100\ \mu m^3$ of water is used to simulate the biological tissue and the environmental temperature is set to be 22 °C. Figure 4(a) shows the configuration of the simulated focal volume of single-plane imaging with a heat power of 4 mW applied. Figure 4(b) shows the configuration of dual-plane imaging where the two focal points are separated by 80 µm and a heat power of 2 mW is applied to each focal volume. The simulated temperature distribution across the section of the simulated volume under the settings of single- and dual-plane imaging is shown in Figs. 4(c) and 4(d), respectively. With the same total heat power, the highest temperature for single- and dual-plane imaging is ~57 °C and ~40 °C, respectively. Therefore, by imaging from two planes separated by 80 µm, the focal point temperature rise has been reduced by ~50%, which can dramatically reduce tissue damage. This is critical in MPM where a large amount of laser power is needed in order to excite 2PM and 3PM signals.

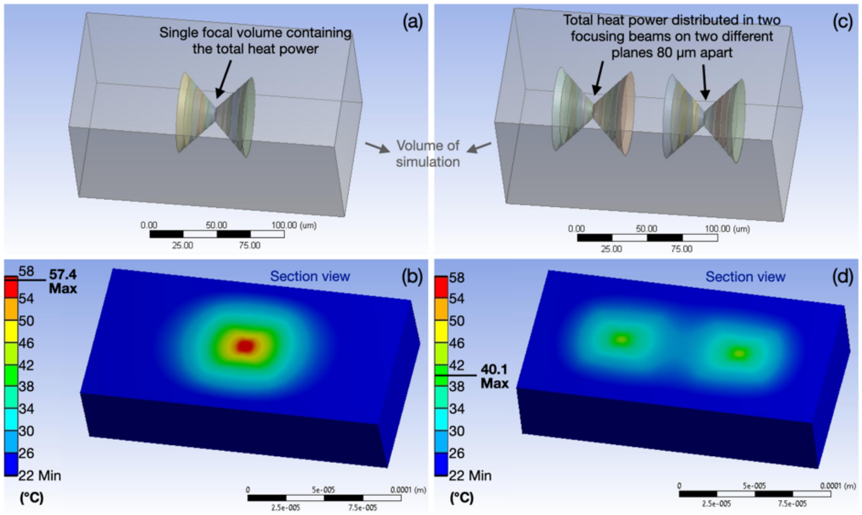
(a)
Single focal volume containing the total heat power
(c)
Total heat power distributed in two focusing beams on two different planes 80 μm apart
Volume of simulation
0.00
50.00
100.00 (um)
25.00
75.00
(b)
Section view
58
57.4 Max
54
50
46
42
38
34
30
26
22 Min
(°C)
0
5e-005
0.0001 (m)
2.5e-005
7.5e-005
(d)
40.1 Max

Fig. 4. Configuration of thermal analysis and obtained temperature distribution of single vs. dual focal-plane imaging under the same total heat power. (a) and (b) are for single focal-plane imaging; (c) and (d) are for dual focal-plane imaging. The total heat power is 4 mW and the environment temperature is set to be 22 °C.

4. Kernel-based nonlinear scaling denoising algorithm

To address the challenge of the large variation in the SNR among the multimodal images and to visualize weak-signal channels in merged multimodal images, we propose a kernel-based nonlinear scaling (KNS) denoising algorithm. Some preliminary results of the KNS algorithm were reported in our previous publication Ref. [16]. Here we will explain the KNS algorithm in detail, and provide improved results and more comprehensive comparisons with various denoising approaches.

The principle of this denoising algorithm is to adjust pixel brightness (intensity) based on a comparison between the global average brightness of the image and the local average brightness of the targeted pixel's surrounding kernel. Based on the main tissue features (e.g., epidermis cells, osteocytes, Volkmann's canals) shown in the image, a square-shaped kernel size of $n_K \times n_K$ pixels is empirically chosen to be ~50% of the size of the tissue feature. For each pixel (x, y) on an image with a dimension of $X \times Y$ pixels, a kernel centered on the pixel (x, y) with a dimension of $n_K \times n_K$ is then generated. The relationship between the scaled brightness I_s and the original brightness I_o of pixel (x, y) is defined as

$$I_s(x, y) = \frac{I_o(x, y)}{R(x, y)}, \tag{1}$$

where $R\,(x, y)$ is a scaling ratio that depends on both the local and global average brightness. In general, $R\,(x, y)$ is higher when the local average to global average ratio is lower. To determine the exact scaling ratio $R(x, y)$, scattered scaling ratio data are first manually tuned on testing low-signal MPM images to reach an optimum denoising result through trial and error. Afterward, an analytical expression of $R\,(x, y)$ is obtained by exponential curve fitting of the scattered data. For the MPM images we have processed, $R\,(x, y)$ can be obtained using the following formula,

$$R(x, y) = 0.1e^{0.94\frac{I_{\mathrm{GM}}}{I_{\mathrm{KM}(x,y)}}} + 0.75e^{0.59\frac{I_{\mathrm{GM}}}{I_{\mathrm{KM}(x,y)}}} \tag{2}$$

Here I_{GM} is the global mean intensity of the entire image, and $I_{\mathrm{KM}}\,(x, y)$ is the local mean intensity within the kernel centered on the pixel (x, y). I_{GM} and $I_{\mathrm{KM}}\,(x, y)$ are respectively calculated as

$$I_{\mathrm{GM}} = \frac{1}{XY} \sum^{X} \sum^{Y} I_{\mathrm{o}}(x, y), \tag{3}$$

$$I_{\mathrm{KM}}(x, y) = \frac{1}{{n_{\mathrm{K}}}^2} \sum^{n_{\mathrm{K}}} \sum^{n_{\mathrm{K}}} I_{\mathrm{o}}(x_{\mathrm{K}}, y_{\mathrm{K}}), \tag{4}$$

where x_{K} and y_{K} are pixel coordinates within the kernel.

This brightness scaling is performed on every pixel of the image. In case the targeted pixel is close to the edge of the image that causes the kernel (centered on that pixel) with the pre-selected dimension to exceed the image boundary, the actual effective kernel will be cropped by the image boundary to a reduced dimension.

Figure 5 shows representative images of low-SNR denoised by KNS and various commonly used denoising methods. The raw images include two THG images of mouse ear pinnae, an SHG image of the lamina of an Acer rubrum (red maple) leaf, and a THG image of a mouse femur bone. In our previous work [16], we demonstrated a comparison between an early-version KNS (the scaling ratio was based on a lookup table, and KNS has no iteration) and outlier removal using the mouse pinna and maple lamina images. Here in Fig. 5, additional denoising methods such as the Gaussian filter and bounded-variable least squares (BVLS) deconvolution algorithm have been added for comparison [43]. Furthermore, KNS is improved to be based on the empirical analytical scaling function Eq. (2) and be able to operate with multiple iterations. To show a more realistic SNR in generally dark images with few scattered signals, we also updated the definition of signal in SNR calculation from the global average brightness in Ref. [16] to the highest local average brightness. In this study, SNR is calculated by taking the ratio between the highest local average and the lowest local average, where the local averages are calculated for each image block with 20 × 20 pixels. The THG contrast of mouse pinna shows the cartilage layer, where the large (diameter of ~30 µm) honeycomb structures could originate from the interfaces created by chondrocyte lacunae [44]. The SHG contrast in the maple lamina could be generated from starch in chloroplasts and cellulose in cell walls [45,46]. The THG image of mouse femur bone shows oval-shaped structures that indicate the location of osteocytes. The THG contrast could originate from the interfaces created by the osteocyte lacunae which are part of the lacuna-canalicular network. All the images contain 328 × 512 pixels. For KNS denoising, the kernel size is selected to be 20 × 20 pixels. KNS is operated on Matlab and the other methods are operated on ImageJ.

As shown in Fig. 5, KNS effectively reduces noise on those low-signal images and can result in an SNR increase of more than 10 dB without significant loss of feature details. In comparison, commonly used traditional methods such as outlier removal, Gaussian filter, and BVLS fail to provide a noticeable SNR improvement to those low-signal MPM images. Thus, KNS outperforms those traditional methods. KNS is able to preserve the brightness of tissue features while reducing the background noise. Most tissue features such as the honeycomb structures in the mouse pinna are well-preserved, and the contrast of tissue features such as circular structures on the leaf lamina is increased. KNS can also be applied iteratively. In Fig. 5, KNS results with iteration numbers $N = 1$ and 2 are shown. Iteratively applying KNS may exponentially increase the SNR, as shown in Fig. 5. However, increasing the iteration number can result in a loss of image details.

Figure 6 compares the denoising results obtained by KNS (with iteration number $N = 1$ or 2) and N2 V. N2 V is a self-supervised deep-learning method that can be trained by a single image without ground truth [38,39]. N2 V training and prediction are operated using the CSBDeep-N2 V plugin of ImageJ [38,47]. The N2 V model is trained by one MPM image and then applied to three other low-signal MPM images as well as the training image itself. In N2 V training, the number of epochs, number of steps per epoch, batch size per step, patch shape, and neighborhood radius are set to be 300, 200, 56, 64, and 5, respectively. The KNS method is applied to the

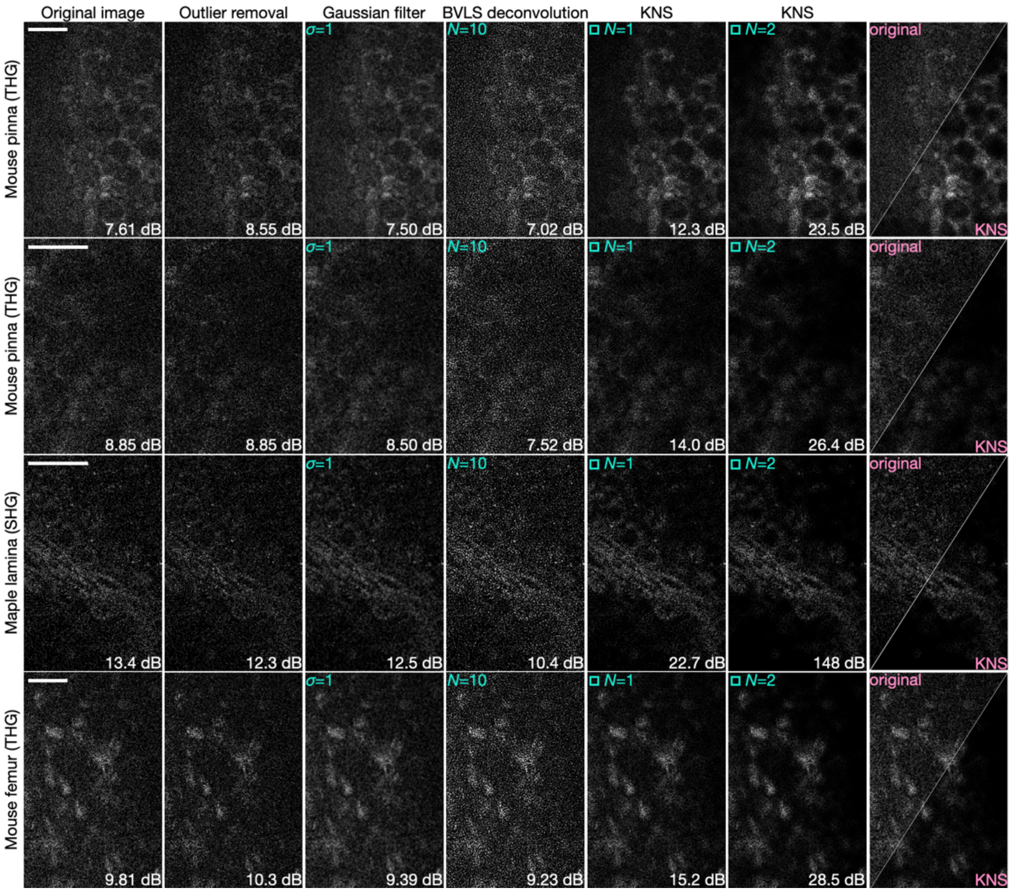
Original image
Outlier removal
Gaussian filter
BVLS deconvolution
KNS
KNS
Mouse pinna (THG)
σ=1
N=10
N=1
N=2
original
7.61 dB
8.55 dB
7.50 dB
7.02 dB
12.3 dB
23.5 dB
KNS
Mouse pinna (THG)
σ=1
N=10
N=1
N=2
original
8.85 dB
8.85 dB
8.50 dB
7.52 dB
14.0 dB
26.4 dB
KNS
Maple lamina (SHG)
σ=1
N=10
N=1
N=2
original
13.4 dB
12.3 dB
12.5 dB
10.4 dB
22.7 dB
148 dB
KNS
Mouse femur (THG)
σ=1
N=10
N=1
N=2
original
9.81 dB
10.3 dB
9.39 dB
9.23 dB
15.2 dB
28.5 dB
KNS

Fig. 5. Comparison of MPM images before and after denoising by outlier removal, Gaussian filter, BVLS deconvolution and KNS. SNR is calculated and shown in each image. σ – radius of the Gaussian filter; N – numbers of iteration. Cyan squares on KNS denoised images indicate the size of the kernel. Scale bar is 40 µm.

same images for comparison. The N2 V model in Fig. 6(a) is trained by a human bone 2PEF image with a relatively high signal level. The image shows osteocytes and their processes with fluorescence from endogenous fluorophores like NADH in the cytoplasm. The denoising result shows that except for the training image, KNS outperforms N2 V in all three low-signal testing images. On the human bone image (also the training image), N2 V shows better results as certain details including some canaliculi are dimmed or lost in the KNS results.

In Fig. 6(b), the N2 V model is trained by a mouse pinna THG image which has an ultra-low signal level. The training does not show a typical successful process as it can be seen that while the training loss decreases through the training process, the validation loss increases. As a result, models from the last epoch and from the lowest validation loss show significantly different performances, and the last-epoch model fails to denoise images. The lowest-loss model is able to preserve some features but also introduces bright background artifacts. For all four images, KNS outperforms the N2 V model in removing the noise while preserving tissue features. On the computation side, both methods ran on a computer equipped with an Intel i9 11900F CPU and an Nvidia RTX 3080 GPU. The N2 V model costed ~21.5 hours to train and took ~1 s for each prediction; KNS took ~34 ms per iteration for each image. While KNS cannot be generally

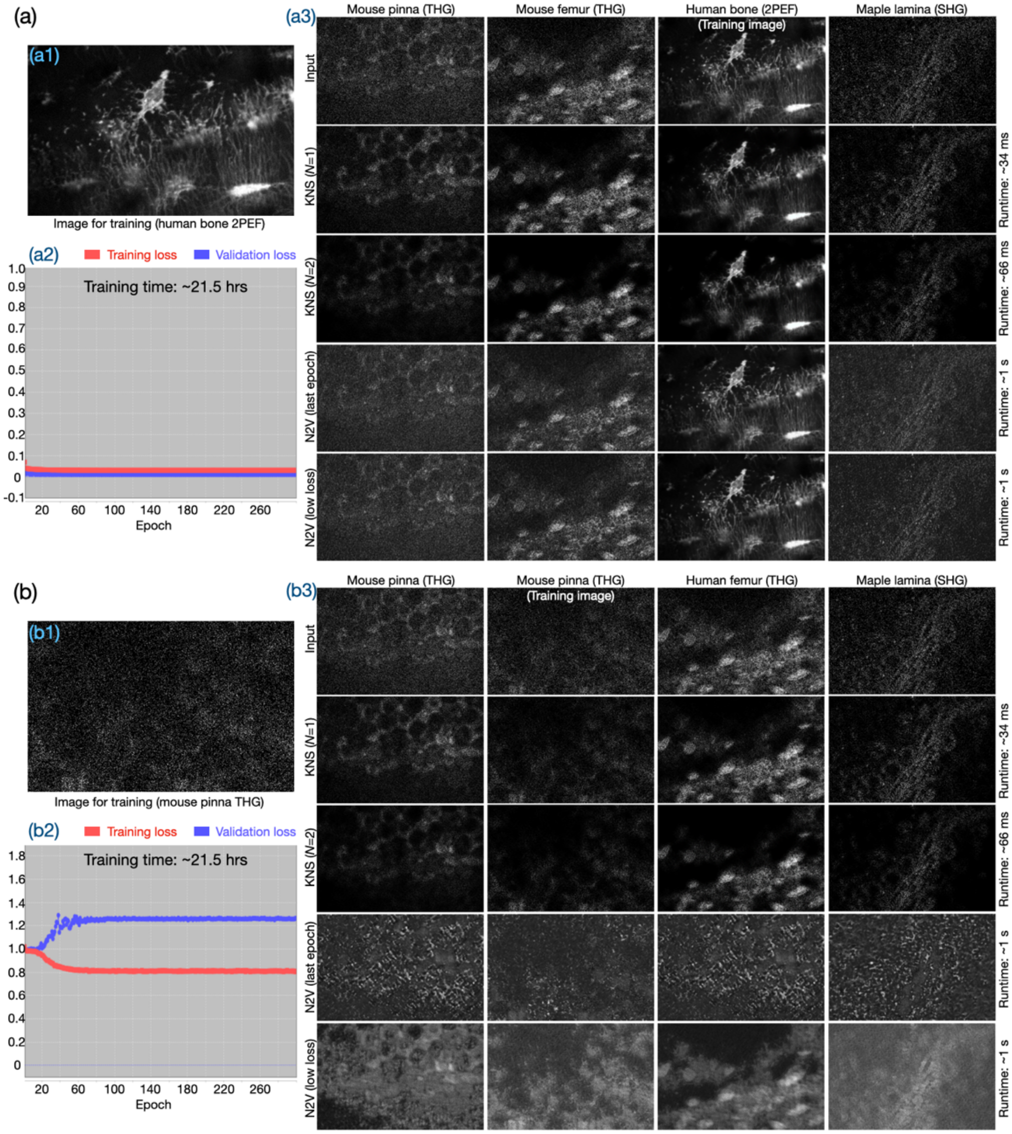
(a)
(a1)
Image for training (human bone 2PEF)
(a2)
Training loss
Validation loss
Training time: ~21.5 hrs
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1
20 60 100 140 180 220 260
Epoch
(a3)
Mouse pinna (THG)
Mouse femur (THG)
Human bone (2PEF) (Training image)
Maple lamina (SHG)
Input
KNS (N=1)
KNS (N=2)
N2V (last epoch)
N2V (low loss)
Runtime: ~34 ms
Runtime: ~66 ms
Runtime: ~1 s
Runtime: ~1 s
(b)
(b1)
Image for training (mouse pinna THG)
(b2)
Training loss
Validation loss
Training time: ~21.5 hrs
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
20 60 100 140 180 220 260
Epoch
(b3)
Mouse pinna (THG)
Mouse pinna (THG) (Training image)
Human femur (THG)
Maple lamina (SHG)
Input
KNS (N=1)
KNS (N=2)
N2V (last epoch)
N2V (low loss)
Runtime: ~34 ms
Runtime: ~66 ms
Runtime: ~1 s
Runtime: ~1 s

Fig. 6. Comparison between KNS and N2 V in processing low-signal noisy images. The N2 V models used in (a) and (b) are trained by (a1) a human bone 2PEF image with a relatively high SNR and (b1) a mouse pinna THG image with an ultra-low SNR, respectively. (a2) and (b2) show the training and validation loss through the training of these two models, respectively.

expected to outperform deep-learning methods such as N2 V, in this comparison, KNS shows its specialty in processing images with ultra-low signal levels. Moreover, its simplicity enables fast computation and makes video-rate real-time processing possible.

To further evaluate the denoising performance, we also compare the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) using simulated data. In this simulation, a real image with high SNR is used as the ground truth (reference image); additive Gaussian noise is added to the ground truth to synthesize a noisy input image; KNS and N2 V methods are then applied to the noisy input image; finally, PSNR and SSIM are computed between the denoised image and the ground truth. Figure 7 shows the results of this comparison. A relatively clean THG image of Amanita muscaria (fly agaric mushroom) spores shown in Fig. 7(a) is used as the ground truth. In Fig. 7(b), Gaussian noise with a mean of zero and a standard deviation of 35 is added to the reference image to synthesize a noisy image which is then inputted to the denoising algorithms. Figure 7(c) shows the KNS denoising result with iteration number $N = 2$. Figures 7(d) and 7(e) are the denoising results using the N2 V model shown in Fig. 6(a) and another N2 V model trained by the input image Fig. 7(b), respectively. Each image is calculated for its PSNR and SSIM with reference to the ground truth. The PSNR and SSIM calculations are performed by the SNR plugin [48] and SSIM index plugin [49] of ImageJ, respectively. The input noisy image has a PSNR and SSIM of 17.51 dB and 0.251, respectively, compared to the ground truth. After denoising by KNS, the PSNR and SSIM improved to 23.54 dB and 0.459, respectively. In contrast, the N2 V processed two results reached only slightly improved PSNRs (18.06 and 18.11 dB) and even deteriorated SSIMs (0.220 and 0.190). In addition, the two N2 V models resulted in observably different results. Compared with the model trained by the bone sample with a good SNR (shown in Fig. 6(a)), the model trained by the input image (Fig. 7(b)) gives a much-blurred prediction (Fig. 7(e) compared with Fig. 7(d)). This shows that because of the black-box effect of machine-learning-based methods, they may generate inconsistent and unpredictable results due to different training and/or input data.

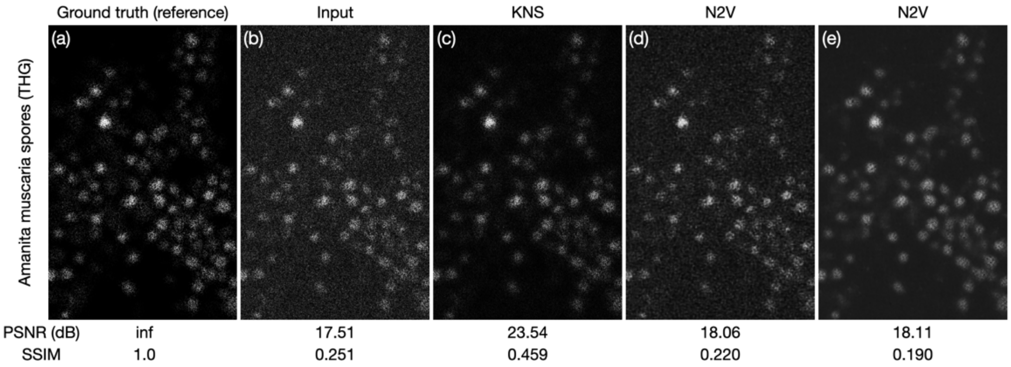

Fig. 7. Comparison of PSNR and SSIM obtained by KNS and N2 V. (a) Ground truth (reference image). (b) Noisy input image created by adding Gaussian noise with a standard deviation of 35 onto (a). (c) KNS (iteration $N = 2$) denoised result. (d) Denoised result by an N2 V model trained by a relatively high SNR image (shown in Fig. 6(a)). (e) Denoised result by an N2 V model trained by the noisy input image (b) with relatively low SNR.

Denoising is particularly important in multimodal imaging. The multiple contrast images may have very different SNRs which can make the merged image very noisy or the weak channel(s) hard to distinguish. Figure 8 shows the contribution of the KNS denoising algorithm to the merging of multimodal images. A mouse pinna and a mouse femur are imaged by 2PEF, SHG, and THG. The THG signal is much weaker than the 2PEF and SHG signals. Without denoising, the noise level is close to the signal level on the THG images, and a large amount of noise is introduced to the merged images, affecting the readability of the other channels. In comparison, the KNS denoising algorithm effectively reduced the noise in the THG images. The merged images show significant improvement with tissue features well preserved and presented.

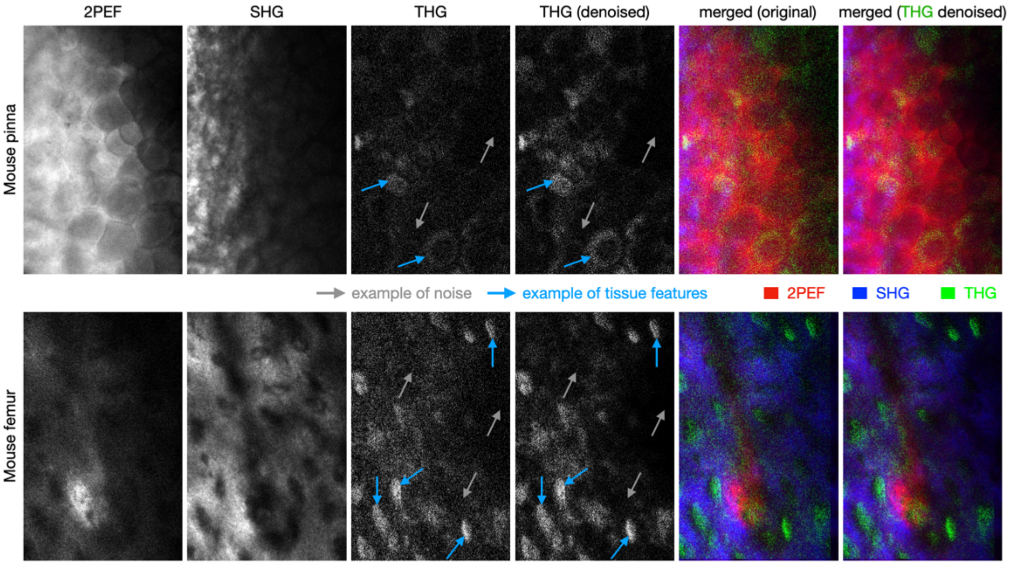

Fig. 8. Comparison of merging multimodal MPM images with and without KNS denoising on the low-signal THG channel. The gray and blue arrows indicate examples of noise and tissue features, respectively.

5. Combined 2PM and 3PM imaging of biological samples

Multimodal MPM with simultaneously acquired 2PM and 3PM images are demonstrated in Fig. 9 on biological samples. The 2PM (2PEF and SHG) and 3PM (THG) signals are acquired simultaneously and label-free, enabling a comprehensive assessment of intricate microstructures while minimizing motion artifacts. Due to the focal shift, the THG image stack is shifted by 80 µm when merging the three channels.

Figure 9(a) shows multimodal MPM images of a fresh yew leaf. The SHG and THG channels show low-level signals and are denoised by KNS (iteration $N = 1$). Cell walls can be clearly seen as bright stripes in 2PEF on the leaf surface and dark stripes beneath the leaf surface. Sub-cellular structures can be identified as dark circular features on the images from a deeper depth (>20 µm). THG contrast shows chloroplasts inside leaf cells underneath the surface.

Figure 9(b) shows multimodal MPM images of a healthy human trabecular bone tissue fixed by formalin and mounted into a wax mold. The THG images are denoised by KNS (iteration $N = 1$). Osteocytes and canaliculi (lacuno-canicular networks) are visualized mainly by THG but also 2PEF contrasts from near the surface up to ~100 µm deep into the tissue. Osteocytes play a vital role in bone homeostasis and mechano-sensation and are shown as oval-shaped structures in THG and 2PEF [50,51]. As THG requires interfaces [20,33], osteocytes may generate THG signals due to their complex morphology and refractive index variations created by the gap of 0.5–1.0 µm between the osteocyte cell body and the lacuna [52,53]. The 2PEF from osteocytes could be generated from endogenous fluorophores like NADH in the cytoplasm. The tiny fibrous structures in the image shown from depths 43, 55, and 74 µm are most likely canaliculi (canals where processes of osteocytes are embedded), which will maintain some extent of their micro-structures even with no metabolism. Given the sensitivity of THG to interfaces [20,33]

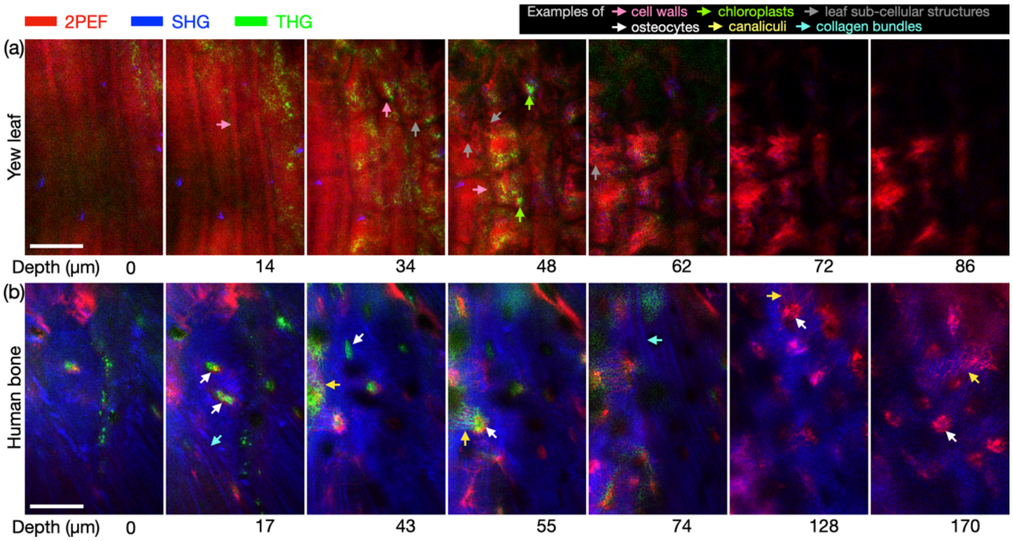
2PEF
SHG
THG
Examples of → cell walls → chloroplasts → leaf sub-cellular structures
→ osteocytes → canaliculi → collagen bundles
(a)
Yew leaf
Depth (μm) 0 14 34 48 62 72 86
(b)
Human bone
Depth (μm) 0 17 43 55 74 128 170

Fig. 9. Simultaneous 2PM and 3PM imaging of biological samples. Red – 2PEF; blue – SHG; green – THG. Scale bar is 40 μm.

and the role of canaliculi as essential communication channels for osteocytes, it is plausible that the observed THG signals originate from interfaces in the intricate networks of those cellular processes. The THG channel does not have sufficient SNR to image deeper beyond ~100 μm while the 2PEF channel can image up to the depth of ~170 μm. This is because three-photon nonlinear effects are generally weaker than two-photon nonlinear effects and the focal spot at a longer excitation wavelength (1580 nm, for 3PM) is not as tight as that at a shorter excitation wavelength (790 nm, for 2PM). The 2PEF contrast also visualizes endogenous fluorophores in other biomolecules and structures such as cartilage within the bone matrix. The SHG contrast delineates the collagen bundles in different depths, showing relatively regular distribution and organization, which confirms the healthy status of the sample.

6. Discussion and conclusions

In this study, a compact multimodal MPM system has been demonstrated to perform 2PM and 3PM imaging simultaneously. The dual-wavelength excitation enables the acquisition of multiple contrasts including 2PEF, SHG, and THG all label-free, which can obtain complementary information about tissues. To achieve simultaneous acquisition of 2PM and 3PM signals, several challenges are addressed. The 2PEF, SHG, and THG channels are separated spectrally by properly selected dichroic beam splitters. The challenge of partial wavelength overlapping between 2PEF and THG has been discussed and a signal-subtraction method is proposed to remove the leaked signal from images. The benefit of reducing motion artifacts and mechanical drift by simultaneous acquisition of multiple channels has been demonstrated. While the 2PM and 3PM image stacks are acquired simultaneously in our study, there is a focal shift of ~80 μm between the 2PM and 3PM stacks. Thus, there still exists a time delay between the corresponding 2PM and 3PM frames from the same tissue depth. This time delay depends on the focal shift, depth scanning step size, and frame rate. To further reduce this time delay and its impact on motion artifacts during *in-vivo* imaging, several approaches can be carried out in future work. First, a faster scanning speed of up to ~4 fps can be applied to reduce the time delay by 10 times. Second, the focal shift can be engineered and reduced by an achromatic design of the objective. Moreover, the focal shift can also be controlled if separate beam paths for the 790 and 1580 nm wavelengths are implemented. Nevertheless, this will increase the complexity and dimension of the system and thus is not ideal for the compact system. On the other side, the focal shift between 790 and 1580 nm can potentially be utilized to acquire images from two planes simultaneously. In addition, since the two beams are focused at two different depths, it can bring about a higher endurance of the total laser power compared to when the two wavelengths are focused on the same plane. When the focal shift is reduced, the laser-induced temperature rise in tissue can increase. Based on our thermal simulation, a focal shift of ∼20–40 µm could be a good balance.

To improve the image quality of low-signal channels and help form clean merged images, the KNS denoising method has been developed. With kernel size chosen by tissue feature dimensions, a kernel is generated centered on each pixel on an image, and the brightness of the central pixel is scaled by a nonlinear equation with variables based on the global and local (kernel) averages. The denoising method is generally universal to different tissue features and has been demonstrated on low-signal THG and SHG images acquired from animal and plant tissue samples. Comparisons between various denoising methods including outlier removal, Gaussian filter, deconvolution, and N2 V have been conducted to verify the denoising performance. KNS outperforms the above-mentioned denoising methods on low-signal MPM images. The reason for KNS to work is that the method judges a pixel's possibility to be noise by its surrounding brightness rather than the pixel itself. This idea provides KNS the capability to distinguish signal and noise from images with signal intensities that are so low that even the part showing signals consists of highly sparkled pixels. As a result, KNS is particularly effective in processing ultra-low signal MPM images. However, for input images already with a relatively high SNR, KNS may cause artifacts that dim some signals. As a scaling method, KNS can alter the relative brightness between different tissue features shown on an image and may not be suitable for quantitative analyses. This effect can be more obvious on bright images with most parts showing signals rather than background noise. With the increased number of iterations, KNS may also cause a loss of image details, saturations, and gridded artifacts. Overall, in the merging of MPM images consisting of multiple channels with highly inconsistent SNRs, KNS improves the image quality by significantly reducing the excessive amount of noise introduced to the superposed image from low-signal noisy channels. KNS provides an effective and fast denoising approach for low-signal images whenever ground truth images are not available.

To demonstrate its capability, depth-resolved multimodal MPM images from biological samples are demonstrated with 2PEF, SHG, and THG signals acquired simultaneously by our compact system. The multimodal imaging approach offers a comprehensive view of microstructures in tissues, facilitating a deeper understanding of tissue composition and potential cellular interactions. Our technique of simultaneous acquisition of 2PM and 3PM signals shows great potential in improving MPM imaging by accelerating the imaging process and reducing possible image artifacts caused by tissue motion and mechanical drift during sequential acquisition. The proposed KNS denoising method further improves the quality of multimodal images and has the potential to be integrated into real-time clinical imaging systems due to its simplicity in computation.

Funding. Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-05913, CHRP 508405-17); Canadian Institutes of Health Research (CPG-151974).

Acknowledgments. We thank Dr. Felipe Eltit and Dr. Michael Cox for providing the human bone sample. We thank the Centre for Disease Modeling and Animal Care Services at the University of British Columbia for providing the mouse tissue samples. We thank Yue (Ashley) Mao, Xuezhu Kou, and Jianren Wu for auxiliary support.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. N. T. Shaked, S. A. Boppart, L. V. Wang, *et al.*, "Label-free biomedical optical imaging," Nat. Photonics **17**(12), 1031–1041 (2023).

2. S. You, H. Tu, E. J. Chaney, *et al.*, "Intravital imaging by simultaneous label-free autofluorescence-multiharmonic microscopy," Nat. Commun. **9**(1), 2125 (2018).
3. G. Borile, D. Sandrin, A. Filippi, *et al.*, "Label-Free Multiphoton Microscopy: Much More than Fancy Images," Int. J. Mol. Sci. **22**(5), 2657 (2021).
4. L. Li, D. Kang, Z. Huang, *et al.*, "Multimodal multiphoton imaging for label-free monitoring of early gastric cancer," BMC Cancer **19**(1), 295 (2019).
5. B. Weigelin, G. Bakker, and P. Friedl, "Third harmonic generation microscopy of cells and tissue organization," J. Cell Sci. **129**(2), 245–255 (2016).
6. T. Sehm, O. Uckermann, R. Galli, *et al.*, "Label-free multiphoton microscopy as a tool to investigate alterations of cerebral aneurysms," Sci. Rep. **10**(1), 12359 (2020).
7. A. Dilipkumar, A. Al-Shemmary, L. Kreiß, *et al.*, "Label-Free Multiphoton Endomicroscopy for Minimally Invasive In Vivo Imaging," Adv. Sci. **6**(8), 1801735 (2019).
8. Y. H. Mohammed, D. S. Barkauskas, A. Holmes, *et al.*, "Noninvasive in vivo human multiphoton microscopy: a key method in proving nanoparticulate zinc oxide sunscreen safety," J. Biomed. Opt. **25**(01), 1 (2020).
9. E. Gavgiotaki, G. Filippidis, M. Kalognomou, *et al.*, "Third Harmonic Generation microscopy as a reliable diagnostic tool for evaluating lipid body modification during cell activation: the example of BV-2 microglia cells," J. Struct. Biol. **189**(2), 105–113 (2015).
10. V. Andresen, S. Alexander, W.-M. Heupel, *et al.*, "Infrared multiphoton microscopy: Subcellular-resolved deep tissue imaging," Curr. Opin. Biotechnol. **20**(1), 54–62 (2009).
11. P. J. Campagnola and L. M. Loew, "Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms," Nat. Biotechnol. **21**(11), 1356–1360 (2003).
12. P. J. Campagnola, A. Lewis, and L. M. Loew, "High-resolution nonlinear optical imaging of live cells by second harmonic generation," Biophys. J. **77**(6), 3341–3349 (1999).
13. P. T. C. So, C. Y. Dong, B. R. Masters, *et al.*, "Two-photon excitation fluorescence microscopy," Ann. Rev. Biomed. Eng. **2**(1), 399–429 (2000).
14. K. Harpel, R. D. Baker, B. Amirsolaimani, *et al.*, "Imaging of targeted lipid microbubbles to detect cancer cells using third harmonic generation microscopy," Biomed. Opt. Express **7**(7), 2849–2860 (2016).
15. K. König, "Multiphoton microscopy in life sciences," J. Microsc. **200**(2), 83–104 (2000).
16. W. Wu, Q. Liu, C. Brandt, *et al.*, "Dual-wavelength multimodal multiphoton microscope with SMA-based depth scanning," Biomed. Opt. Express **13**(5), 2754–2771 (2022).
17. L. Huang, X. Zhou, Q. Liu, *et al.*, "Miniaturized multimodal multiphoton microscope for simultaneous two-photon and three-photon imaging with a dual-wavelength Er-doped fiber laser," Biomed. Opt. Express **11**(2), 624–635 (2020).
18. N. Olivier, M. A. Luengo-Oroz, L. Duloquin, *et al.*, "Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy," Science **329**(5994), 967–971 (2010).
19. S. Mehravar, B. Banerjee, H. Chatrath, *et al.*, "Label-free multi-photon imaging of dysplasia in Barrett's esophagus," Biomed. Opt. Express **7**(1), 148–157 (2016).
20. W. Wu and S. Tang, "Harmonic generation and impact of phase matching in multimodal multiphoton microscopy," IEEE J. Sel. Top. Quantum Electron. **29**(4: Biophotonics), 1–9 (2023).
21. Q. Cheng, H. Chung, R. Schubert, *et al.*, "Protein-crystal detection with a compact multimodal multiphoton microscope," Commun. Biol. **3**(1), 569 (2020).
22. F. Akhoundi, Y. Qin, N. Peyghambarian, *et al.*, "Compact fiber-based multi-photon endoscope working at 1700nm," Biomed. Opt. Express **9**(5), 2326–2335 (2018).
23. C. Lefort, "A review of biomedical multiphoton microscopy and its laser sources," J. Phys. D: Appl. Phys. **50**(42), 423001 (2017).
24. W. R. Zipfel, R. M. Williams, R. Christie, *et al.*, "Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation," Proc. Natl. Acad. Sci. U. S. A. **100**(12), 7075–7080 (2003).
25. K. Kieu, S. Mehravar, R. Gowda, *et al.*, "Label-free multi-photon imaging using a compact femtosecond fiber laser mode-locked by carbon nanotube saturable absorber," Biomed. Opt. Express **4**(10), 2187–2195 (2013).
26. D. Septier, V. Mytskaniuk, R. Habert, *et al.*, "Label-free highly multimodal nonlinear endoscope," Opt. Express **30**(14), 25020–25033 (2022).
27. F. Aptel, N. Olivier, A. D. Besseau, *et al.*, "Multimodal Nonlinear Imaging of the Human Cornea," Invest. Ophthalmol. Visual Sci. **51**(5), 2459–2465 (2010).
28. W. Wu, Q. Liu, C. Brandt, *et al.*, "Multimodal Multiphoton Microscope with Depth Scanning," Proc. SPIE **11965**, 119650J (2022).
29. A. Filippi, E. D. Sasso, L. Iop, *et al.*, "Multimodal label-free *ex vivo* imaging using a dual-wavelength microscope with axial chromatic aberration compensation," J. Biomed. Opt. **23**(9), 091403 (2018).
30. D. G. Ouzounov, T. Wang, M. Wang, *et al.*, "In vivo three-photon imaging of activity of GCaMP6-labeled neurons deep in intact mouse brain," Nat. Methods **14**(4), 388–390 (2017).
31. S. Weisenburger, F. Tejera, J. Demas, *et al.*, "Volumetric Ca^{2+} Imaging in the Mouse Brain Using Hybrid Multiplexed Sculpted Light Microscopy," Cell **177**(4), 1050–1066.e14 (2019).
32. E. A. Gibson, O. Masihzadeh, T. C. Lei, *et al.*, "Multiphoton Microscopy for Ophthalmic Imaging," J. Ophthalmol. **2011**, 870879 (2011).

33. Y. Barad, H. Eisenberg, M. Horowitz, *et al.*, "Nonlinear scanning laser microscopy by third harmonic generation," Appl. Phys. Lett. **70**(8), 922–924 (1997).
34. W. Wu and S. Tang, "Multimodal MPM Image Processing: Nonlinear Denoise Method Based on Kernel Size Determined by Tissue Feature," in *Biophotonics Congress: Optics in the Life Sciences* (2023), paper NTu1C.6.
35. S. Niu, L. Guo, Y. Li, *et al.*, "Boundary-Preserved Deep Denoising of Stochastic Resonance Enhanced Multiphoton Images," IEEE J. Transl. Eng. Health Med. **10**, 1–12 (2022).
36. F. Chollet, *Deep Learning with Python*, 2nd ed. (Manning Publications, 2021).
37. A. Jaiswal, A. R. Babu, M. Z. Zadeh, *et al.*, "A Survey on Contrastive Self-Supervised Learning," Technologies **9**(1), 2 (2020).
38. A. Krull, T.-O. Buchholz, and F. Jug, "Noise2Void - Learning Denoising from Single Noisy Images," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 2124–2132.
39. J. Lehtinen, J. Munkberg, J. Hasselgren, *et al.*, "Noise2Noise: Learning image restoration without clean data," in *International Conference on Machine Learning (ICML)* (2018), pp. 2965–2974.
40. Z. C. Lipton, "The Doctor Just Won't Accept That!," arXiv, arXiv:1711.08037v2 (2017).
41. C. Brandt, W. Wu, Q. Liu, *et al.*, "Endoscopic MPM objective designed for depth scanning," Proc. SPIE **11937**, 1193706 (2022).
42. W. Wu, C. Brandt, and S. Tang, "Simultaneous 2-photon and 3-photon multimodal signal acquisition for multiphoton microscopy," Proc. SPIE **12384**, 123840H (2023).
43. D. Sage, L. Donati, F. Soulez, *et al.*, "DeconvolutionLab2: An open-source software for deconvolution microscopy," Methods **115**, 28–41 (2017).
44. P. Sosnowski, P. Sass, P. Słonimska, *et al.*, "Regenerative Drug Discovery Using Ear Pinna Punch Wound Model in Mice," Pharmaceuticals **15**(5), 610 (2022).
45. G. Cox, N. Moreno, and J. Feijo, "Second-harmonic imaging of plant polysaccharides," J. Biomed. Opt. **10**(2), 024013 (2005).
46. S. Chu, I. Chen, T. Liu, *et al.*, "Multimodal nonlinear spectral microscopy based on a femtosecond Cr:forsterite laser," Opt. Lett. **26**(23), 1909–1911 (2001).
47. ImageJ Wiki, "N2 V," https://imagej.net/plugins/n2v#publication-noise2void—learning-denoising-from-single-noisy-images.
48. D. Sage, "SNR, PSNR, RMSE, MAE - ImageJ's plugin to assess the quality of images," http://bigwww.epfl.ch/sage/soft/snr/
49. G. P. Renieblas, "Calculate SSIM Index," https://imagej.nih.gov/ij/plugins/ssim-index.html
50. D. Tokarz, R. Cisek, M. N. Wein, *et al.*, "Intravital imaging of osteocytes in mouse calvaria using third harmonic generation microscopy," PLoS One **12**(10), e0186846 (2017).
51. R. Genthialetal, M. Gerbaix, D. Farlay, *et al.*, "Third harmonic generation imaging and analysis of the effect of low gravity on the lacuno-canalicular network of mouse bone," PLoS One **14**(1), e0209079 (2019).
52. R. Florencio-Silva, G. R. Sasso, E. Sasso-Cerri, *et al.*, "Biology of bone tissue: Structure, function, and factors that influ- ence bone cells," Biomed. Res. Int. **2015**, 421746 (2015).
53. L. You, S. Weinbaum, S. C. Cowin, *et al.*, "Ultrastructure of the osteocyte process and its pericellular matrix," Anat. Rec. Part A: Discoveries Mol. Cell Evol. Biol. **278A**(2), 505–513 (2004).